Remarks

Claims 1-5 and 7-11 are pending in the application. Claim 11 is allowed. Claims 1-5 and 7-9 are rejected. Claim 10 is objected to.

Applicants take note that the previous rejections based on Toklu et al. (U.S. Patent No. 6,549,643) are considered moot, and the current Final Rejection is based entirely on new grounds of rejection, which was available to the Examiner in the prior rejection.

Claims 1-5 and 7-9 are rejected under 35 USC 103(a) as being unpatentable over Wee et al. (U.S. Patent No. 6,104,441) in view of Park et al. (U.S. Patent No. 6,597,738).

The Examiner states that Wee discloses a method for summarizing a video in Figures 6-8, and at column 5, lines 50-60. With all due respect, this is completely incorrect. Instead, Wee states, at column 4, lines 50-51, that "FIG. 6 is a software block diagram indicating actions in converting a P frame to an I frame" and, at column 11, lines 35-36, that "[c]onversion of a P frame 241 to an I frame 243 is indicated in block diagram form by FIG. 6."

Converting a P frame to an I frame does not lessen the number of frames in the video. One of ordinary skill in the art would readily understand that an operation that does not diminish the number of frames in a video can never summarize a video. The Applicant fails to understand how the Examiner can

equate a simple frame conversion process with a summarization process as claimed.

Similarly, Wee states, at column 4, lines 52-54, that "FIG. 7 is a software block diagram indicating actions in converting a B frame to a B' frame having only forward or backward dependencies, but not both" and, at column 12, lines 8-12, that "[c]onversion of a B frame to a B' frame is illustrated using FIG. 7, and the conversion employs a subroutine 271 which in implementation is similar to that described above in connection with FIG. 6."

This is also nothing but a simple conversion process and not a summarization process as claimed.

Similarly, Wee states, at column 4, lines 55-56, that "FIG. 8 is a software block diagram indicating actions in converting an I frame to a P frame" and, at column 12, lines 59-60, that "[a]s seen in FIG. 8, an I frame 301 is processed using a second subroutine 303 to yield a P frame 305. "

This is yet another conversion, and not a summarization.

At column 5, lines 50-60:

> "For example, the two reference blocks 117 and 119 in the preferred embodiment indicate that each set of frames is converted (without performing an inverse discrete cosine transform "IDCT") to a compressed data subset 121 and 123, representing only those frames that are to be

> retained. These subsets are then combined, as indicated by a summing junction 125, and rate control is then applied to prevent data underflow or overflow, as indicated by reference block 127. The editor's finished product is then provided in a fully MPEG compliant compressed image output sequence 129, whereby editing has been accomplished to produce a composite image sequence 131, without fully decoding each of the two image sequences 103 and 105 to the image domain."

Wee makes it abundantly clear that what he is doing is converting one frame type to another frame type.

Those of ordinary skill in the art would confuse the conversion described by Wee with the summarization claimed.

From the Merriam-Webster Dictionary:

Summarize: to reduce to a summary

Convert: to change from one form to another

The Examiner goes on to state that Wee describes measuring a cumulative motion activity at column 9, lines 30-35:

> "The cutting and splicing together of two image sequences is better illustrated by FIG. 4. It will be assumed that these image sequences have been dequantized and expanded to motion vector- ("MV-") and discrete cosine transform- ("DCT-") based representations. First, the editor proceeds to form a head data stream by identifying any cut points in a first sequence of frames, as identified by a processing block 201 of FIG. 4."

Applicants cannot find anything about measuring motion activity in Wee. All that Wee does is an inverse quantization to recover motion vectors from the compressed domain frames. To measure motion activity, as defined in the application, one must do some measurement on the motion vectors: "One measure of motion activity can be the average and variance of the magnitude of the motion vectors, see Peker et al. '*Automatic measurement of intensity of motion activity*,' Proceedings of SPIE Conference on Storage and Retrieval for Media Databases, January 2001. However, there are many variations possible, depending on the application." see page 5, lines 12-16. Wee only decodes and encodes motion vectors.

The Examiner states that Wee selects key frames from the compressed video according to the cumulative motion activity intensity at column 8, lines 35-53. With all due respect this is completely incorrect again. There, Wee states:

> "However, contrary to conventional wisdom, the preferred editor in this example performs processing without converting each of the two image sequences to the image domain prior to performing editing. In particular, the first image sequence is first processed by converting frame B_5 to remove its dependencies upon any subsequent frames, e.g., frame P_7 The editor, via a "transcoding" process described below, determines that for each block of frame B_5 that depends upon a block of frame P_4 (and P_7) , it should add an estimate of 1/2 the DCT coefficients for the block identified by motion vectors pointing to frame P_4, while subtracting 1/2 of the DCT coefficients for the block identified by motion vectors pointing to frame P_7, and removes dependencies attached by frame B_5 to future frames. The

> editor, also identifying that frame P_G depends upon frames P_E, P_C and I_A, uses transcoding to convert each DCT block of frame P_G to reform frame P.sub.G into an I frame, e.g., I_G, by first converting frame P_C to an independent frame, I_C, then converting frame P_E to an independent frame, I_E Represented in transmission order, each of the first and second sequences of images then appears as follows."

Measuring activity, as claimned, and using the motion activity to select key frames is not described by Wee, **anywhere**.

The Examiner states that Wee concatenates key frames selected according to motion activity in a temporal order to form a summary at column 9, lines 1-25. Applicants find **nothing** in the reference that would indicate that Wee concatenates key frames selected on a basis of cumulative motion activity. With all due respect, the Examiner is incorrect on all three claimed steps of measuring, selecting and concatenating. Wee does not have any of the claimed limitations.

The Examiner states that Wee fails to disclose using an *average motion vector activity*. This is not what is claimed. What is claimed is "measuring a cumulative motion activity intensity." First, an *average* is not an *accumulation*. Second, a *motion vector* is not *motion activity*.

Then the Examiner states that Park "discloses the use of an *average motion vector* magnitude." Again, an average is not an accumulation, and a motion vector magnitude is not motion activity. Furthermore, Park does not disclose "dominant motion activity intensity." The only thing that Park accumulates

is a motion intensity histogram. Park, like Wee, has nothing to do with video summarization as claimed. Combining Park with Wee does not cure the defects of Wee, since neither reference summarizes.

Claimed is a partitioning of a compressed video into a plurality of segments, each segment having a substantially equal amount of cumulative motion activity intensity; and selecting one key-frame from each segment. Applicants cannot find anything in Wee that indicates a segmentation based on cumulative motion activity intensity or selecting key frames from each segment. From Wee at column 8:

> frame. Consequently, the transmitted (or coded) order for the sequence indicated above would be as follows;
>
> $I_1P_4B_2B_3P_7B_5B_6I_{10}B_8B_9$.
>
> 1. Image Splicing.
>
> In a first hypothetical, the image sequence
>
> $I_1B_2B_3P_4B_5B_6P_7B_8B_9I_{10}$
>
> is to be cut and spliced after frame 5 to frame 7 of a second image sequence, identified first in temporal order and then in transmission order as follows;
>
> $I_AB_BP_CB_DP_EB_FP_GB_HB_II_J$,
>
> $I_AP_CB_BP_EB_DP_GB_FI_JB_HB_I$.
>
> That is to say, one desires to form a third image sequence with frames identified as follows;
>
> $\underline{I_1}P_4\underline{B_2}\underline{B_3}P_7\underline{B_5}B_6I_{10}B_8B_9$–$I_AP_CB_BP_E\underline{B_D}\underline{P_G}B_FI_J\underline{B_H}\underline{B_I}$
>
> where frames appearing in bold, italicized, underlined font are to be kept, with the remaining frames being discarded.

As state above neither Wee nor Park have anything to do with segments that are one less than a number of desired key-frames in the summary, selecting a first frame of each segment as one of the key-frames of the summary, or selecting a last frame of the compressed video as a last key-frame of the summary. In the rejection, the Examiner has not addressed every recited limitation. MPEP 707.07(f) further mandates that "where a major technical

rejection is proper, it should be stated with a full development of the reasons rather than by a mere conclusion coupled with some stereotyped expression." The rejection by the Examiner is a mere conclusion, without a full development of reasons.

There is no selecting a middle frame according to the accumulated motion activity intensity of each segment as one of the key-frames at column 8:

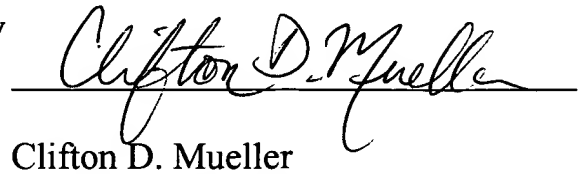
> vectors pointing to frame P_7, and removes dependencies attached by frame B_5 to future frames. The editor, also identifying that frame P_G depends upon frames P_E, P_C and I_A, uses transcoding to convert each DCT block of frame P_G to reform frame P_G into an I frame, e.g., I_G, by first converting frame P_C to an independent frame, I_C, then converting frame P_E to an independent frame, I_E. Represented in transmission order, each of the first and second sequences of images then appears as follows;
>
> $I_1P_4B_2B_3P_7B_5B_6I_{10}B_8$ B_9, and $I_AP_CB_BP_EB_DI_GB_FI_JB_H$ B_I.

At column 7, Park describes a motion direction descriptor, not the claimed motion activity intensity.

Nowhere does Wee describe shot boundaries.

It is believed that this application is now in condition for allowance. A notice to this effect is respectfully requested. Should further questions arise concerning this application, the Examiner is invited to call Applicant's agent at the number listed below. Please charge any shortage in fees due in connection with the filing of this paper to Deposit Account 50-0749.

Respectfully submitted,
Mitsubishi Electric Research Laboratories, Inc.

By 

Clifton D. Mueller
Agent for the Assignee
Reg. No. 57,836

201 Broadway, 8th Floor
Cambridge, MA 02139
Telephone: (617) 621-7517
Customer No. 022199